

2005

Evaluating a Moral Thinking Assessment Model for Evangelical Christian Liberal Arts Colleges

Michael A. Hayes

Lee University

Follow this and additional works at: https://pillars.taylor.edu/acsd_growth



Part of the [Higher Education Commons](#)

Recommended Citation

Hayes, Michael A. (2005) "Evaluating a Moral Thinking Assessment Model for Evangelical Christian Liberal Arts Colleges," *Growth: The Journal of the Association for Christians in Student Development*: Vol. 5 : No. 5 , Article 5.

Available at: https://pillars.taylor.edu/acsd_growth/vol5/iss5/5

This Article is brought to you for free and open access by Pillars at Taylor University. It has been accepted for inclusion in Growth: The Journal of the Association for Christians in Student Development by an authorized editor of Pillars at Taylor University. For more information, please contact pillars@taylor.edu.

Evaluating a Moral Thinking Assessment Model for Evangelical Christian Liberal Arts Colleges

by Michael A. Hayes

Michael A. Hayes Ph.D., is Director of Student Development at Lee University. This research was supported in part by financial assistance from Lee University and the Educational Administration and Policy Studies program at the University of Tennessee, Knoxville. This study is based on a dissertation completed at the University of Tennessee, Knoxville.

Abstract

A model using moral judgment and cultural ideology (political and religious ideology) for predicting moral thinking about critical social and political issues, developed by Narvaez, Getz, Rest, and Thoma (1999), was assessed for utility with students at evangelical Christian liberal arts colleges. Freshmen ($N = 199$) and seniors ($N = 230$) from 2 evangelical Christian liberal arts colleges participated, completing the Defining Issues Test 2, Inventory of Religious Belief, and Attitudes Toward Human Rights Inventory. The regression model predicted a significant amount of variance for the students in this study; however, the R^2 value (.22) was much smaller than in Narvaez et al. (.67). The conclusions from the study were that the model could be used to predict moral thinking for students at these colleges, even though the amount of variance explained by the model was fairly low. Also, the model does not have good statistical fit for students at these colleges, indicating the need for further development of assessment models.

Moral Thinking Assessment Model

In many ways the mission and philosophy of American higher education has changed drastically since the founding of the early American colleges and universities. While many of the early schools focused on training men for the ministry, today the academy is in many ways a “multiversity” (Kerr, 1995), embracing a wider diversity of students, pluralistic values, and purposes (Brubacher & Rudy, 1997). In spite of all of the changes, the development of students’ morality has remained a distinct objective (Evans, Forney, & Guido-DiBrito, 1998; Nucci & Pascarella, 1987). In fact, some (Pascarella, 1997) see that American colleges and universities have a “clearly defined role in developing individuals who can both think and act morally” (p. 47) and serve “as an excellent laboratory for moral development” (Evans et al., 1998, p. 172).

This objective of facilitating students’ ethical and moral development is at the core of the mission of evangelical Christian liberal arts colleges (Holmes, 1991). As Holmes (1987) writes, “In a Christian college one must come to see the distinctive ingredients and bases of Christian values and will, one hopes, make those values one’s own” (p. 32). Moreover, a hallmark of these institutions is their goal of integrating faith, living, and learning (Council for Christian Colleges and Universities, 2000; Holmes, 1987; Peterson’s, 1998), to help students weave together their beliefs and their behaviors

(Garber, 1996). According to Holmes (1991), the Christian college's role in moral development goes beyond indoctrination to helping students learn how to think about issues. This goal is embedded in the broader liberal arts tradition. It is paramount for students to learn to analyze their environments, to think critically about issues, and to make informed decisions based on principles related to their faith, "to be Christian through and through" (Holmes, 1991, p. 8). The focus is on educating students to make decisions about their values rather than making the decisions for them.

As these schools strive to develop students academically and morally, they face a multifaceted challenge in the process. On one hand, they encourage students to think for themselves, particularly as it relates to significant moral and social issues. However, this process is influenced strongly by the religious orientation of the campuses, especially on the more politically and theologically conservative campuses. The conservatism of these schools often is reflected in the campus milieu through behavioral standards set forth and enforced by the institution leading to a potential conflict between encouraging students to critically evaluate issues and behavioral options to reach their own decisions, while concomitantly attempting to shape students' character from a perspective that may lean towards an *in loco parentis* approach by limiting and perhaps dictating their choices. In fact, some posit that students living on such campuses might sacrifice themselves academically while attempting to achieve some sense of moral superiority (McNeel, 1994). Therefore, Christian higher education institutions face a challenge in terms of educating students to think for themselves and encouraging them to critically reflect on their experiences (Dirks, 1988; Holmes, 1991), while providing this education within a conservative Christian environment. Rest, Narvaez, Bebeau, and Thoma (1999) identified this conundrum.

If orthodox religious teachings emphasize the moral authority that is transcendent, supernatural, and beyond attempts at human understanding—and that it is improper and sinful to question, critique, and scrutinize its authority—then orthodoxy may reinforce itself, making difficult movement out of orthodoxy. (p. 121)

Can students in these settings advance in their moral judgment while holding to conservative religious and political ideologies?

A substantial body of literature exists on how colleges influence the moral judgment of their students (McNeel, 1991, 1992; Pascarella, 1997; Pascarella & Terenzini, 1991; Rest, 1986; Rest & Narvaez, 1998b; Rest, Narvaez, Bebeau, et al., 1999). The single best predictor of a person's moral judgment is the amount of formal education completed (Rest, Narvaez, Bebeau, et al., 1999). Therefore, as students progress through their undergraduate experiences, their moral judgment, according to moral judgment models based on Kohlberg's (1981) research, should be developing. However, there is a mixed body of literature on how education at religiously-affiliated influences moral judgment (Beller, Stoll, Burwell, & Cole, 1996; Getz, 1984). Getz (1984) reviewed the findings of the literature on moral judgment and attendance at church-affiliated educational institutions, identifying five studies in this area. In three of the studies the students scored higher than their counterparts in moral judgment, in one study students scored lower, and in the final study there were no significant differences. Although the findings in terms of religious education were mixed, Getz' review (1984) of eight studies that focused on the relationship between moral judgment and religious

ideology or belief showed a more consistent relationship. Seven of the eight studies found that religiously liberal people scored higher in moral judgment, while the eighth study found no significant relationship. Based on these results, she recommended continued research on how dogmatic political and religious ideology relate to moral judgment and on what types of religious education might foster or hinder growth in moral judgment.

One key limitation of the body of literature on moral development and higher education, including Christian higher education, is the primary focus on moral judgment instead of other areas of moral and ethical concern. In light of this and the role and mission of its member institutions, the Council for Christian Colleges and Universities (CCCCU) initiated a six-year (1994-2000) research project entitled, "Taking Values Seriously: Assessing the Mission of Church-Related Higher Education," to determine the extent to which member schools were influencing student values. The results of the project indicated that students at the CCCCU institutions rated themselves as political conservatives more often than their counterparts at Protestant and general four-year colleges on the Cooperative Institutional Research Program (CIRP) survey as freshmen (Baylis, 1997) and on the College Student Survey (CSS) as seniors (Burwell, 1997). However, both CCCCU freshmen and seniors tended to score similar to the Protestant and general four-year college groups on the actual political and social issues items indicating that they may be more politically and socially liberal than had been thought, at least when measured by their stances on specific issues of current social importance. These findings would seem to suggest that Christian liberal arts schools are not fulfilling their missions of influencing their students' values on significant social and political issues in the direction or to the extent that they had purposed. This is problematic given that many of these schools market their superior ability in developing students morally (Beller et al., 1996; Dobson, 1998). Obviously, these institutions need accurate assessment models to measure mission attainment in this area and to validate their claims.

Building on a previous study by Getz (1985) in which she developed a measure of attitudes toward human rights and major social and political issues, Narvaez, Getz, Rest, and Thoma (1999) studied the relationships among moral judgment (using the original Defining Issues Test [DIT]), religious ideology, political ideology, and religious orientation and how they predict attitudes toward human rights and major social issues. They found that political and religious ideology combined into a factor that they called cultural ideology. This, in conjunction with moral judgment, combined to form a variable they called orthodoxy/progressivism, which in turn yielded strong regression coefficients in predicting the participants' moral thinking (i.e., attitudes toward human rights) in a sample drawn from two Protestant churches ($R = .79$; $N = 96$) and in another sample consisting of students from a local state university ($R = .77$; $N = 62$). Individuals who were more progressive tended to score more liberally on their attitudes on human rights, while more orthodox people tended to score more conservatively. Therefore, orthodoxy/progressivism predicted a significant amount of variance in moral thinking on significant social issues.

Rest, Narvaez, Thoma, and Bebeau (1999) replicated the previous study (Narvaez et al., 1999) in an attempt to establish the validity of the second version of the DIT (DIT2). To do so, 200 respondents from four levels of education (ninth-grade students,

senior high graduates, college seniors, and graduate school and professional school students) completed both the DIT and DIT2 and the same measures of religiosity, political ideology, and attitudes toward human rights as used by Narvaez et al. (1999). They found that the multiple regression model with the original DIT as the measure of moral judgment produced a multiple R of .56 ($df = 151$), while the model with the DIT2 produced a multiple R of .58 ($df = 191$). The authors found that their sample scored more conservatively on moral judgment, religious ideology, and attitudes toward human rights as compared to the Narvaez et al. (1999) study. In addition, the participants rated themselves as more politically conservative. Since the R values were somewhat lower in this study with a more conservative sample as compared to the more liberal sample in the Narvaez et al. (1999) study, Rest, Narvaez, Thoma, et al. (1999) recommended additional research to determine whether the strength of the regression model would remain stable between liberal and conservative samples. This current project was undertaken in response to this recommendation, replicating the study with a population with more education and a higher degree of conservatism, and to the need for a valid model for assessing students' moral thinking at Christian colleges and universities.

In addition, the studies by Narvaez et al. (1999) and Rest, Narvaez, Thoma, et al. (1999) asserted that moral judgment and cultural ideology, an unobserved variable comprised of political and religious ideology, combine to "produce moral thinking" (p. 478), thereby claiming causal processes among the variables. Structural equation modeling is used to confirm proposed theories implying causation, particularly with unobserved variables, those which cannot be observed directly. If a model has good statistical fit, "the model argues for the plausibility of postulated relations among variables; if it is inadequate, the tenability of such relations is rejected" (Byrne, 2001, p. 3). Although the model proposed in the Narvaez et al. and Rest, Narvaez, Thoma, et al. studies proffered a causal theory, neither study used structural equation modeling to assess the fit of the model.

The purpose of this study was to assess the utility of a model used to predict moral thinking on major social issues (Narvaez et al., 1999) in evangelical Christian liberal arts institutions. The model used moral judgment and cultural ideology, which was comprised of political ideology and religious ideology, to predict to moral thinking. In addition, this study sought to extend the model by assessing the statistical fit of the model. The research questions that framed this study were:

1. Do moral judgment and cultural ideology (i.e., political ideology and religious ideology) combine to explain a significant amount of the variance in moral thinking in students at evangelical Christian liberal arts colleges and universities as in the Narvaez et al. (1999) study?
2. Does the model predicting moral thinking from moral judgment and cultural ideology for students at evangelical Christian colleges have good statistical fit?

By answering these questions, the study will provide evidence of whether the model used in Narvaez et al. (1999) is generalizable to a very conservative population with higher levels of formal education. Since Christian higher education institutions accent

student moral development, they need to develop ways to assess whether their students do indeed acquire high levels of moral thinking.

Method

Participants

A multistage sampling procedure was used to select students for this project (Babbie, 1990; Fowler, 1993; Henry, 1990). The first stage involved selecting schools that met specified criteria. The schools had to be:

1. evangelical Christian colleges with a holiness tradition
2. fully accredited by the Southern Association of Colleges and Schools (SACS)
3. full members of the CCCU

In selecting the schools, attention was given to using a homogeneous sample to determine whether the predictive variables would still account for a significant amount of variability in moral thinking with this group of students. In addition, the study sought to delimit the schools by instituting undergraduate enrollment size requirements of more than 1,000 to ensure the availability of enough students to participate in the project. Of the 29 CCCU schools accredited by SACS, three schools met the criteria and were invited to participate. Although all three schools initially agreed to participate, only two actually did. More than one school was sampled to assess for any institutional effects as part of a larger research project.

The second stage of sampling involved selecting students at these schools. While this article is focused on the utility of the moral thinking assessment model for Christian liberal arts schools, other research questions were addressed as part of the larger study. One of the questions of the larger study sought to compare how new and advanced students performed on the model; therefore, both freshmen and seniors were sampled from each school. A convenience sampling strategy was utilized by administering the questionnaires to students in classes primarily consisting of first-year students or seniors at the two schools (Henry, 1990). Institutional research personnel at each school generated a list of courses from all departments that were identified as freshman- or senior-focused or were clearly scheduled for students to complete early in the general education core or nearer to the end of their programs of study. Once these lists were generated, course enrollment numbers were examined to ensure adequate sampling. Then, the necessary numbers of courses were selected to ensure a sufficient sample. Research personnel at the schools sought permission from the course instructors and scheduled dates for data collection.

The researcher visited numerous courses at each campus. The schools were given pseudonyms (Epsilon College and Theta College) to protect their confidentiality. At Epsilon College, the researcher visited five introductory psychology courses to administer the battery of instruments to their first-year students and gathered data from eleven upper division courses from a variety of disciplines to collect senior data. In addition, the researcher visited four introductory Bible courses at Theta College to collect data from their freshmen and administered the battery in five upper division courses from five different departments. The total numbers in the sample from Epsilon College and Theta College were 199 and 230 respectively, yielding a total sample size

of 429. At Epsilon College, 94 freshmen participated along with 105 seniors, while 111 freshmen and 119 seniors completed valid protocols from Theta College. For the entire sample 262 (61.1%) were female, and 167 (38.9%) were male. The participants were advised of the nature of the study, were permitted to withdraw from the study at any time without penalty, and completed an informed consent form.

Materials

Each participant was asked to complete three instruments. These included the DIT2 (Rest & Narvaez, 1998a), the Inventory of Religious Beliefs (Brown & Lowe, 1951), and the ATHRI (Getz, 1985). The political ideology item was asked on the DIT2 as part of the standard data collected on that test. The respondents provided other demographic data on that scale as well, specifically educational level, gender, and age.

Defining Issues Test 2 (DIT2). The DIT2, a paper-and-pencil test, was used to measure moral judgment for this study. According to Rest and Narvaez (1998b), the DIT2 is based on Lawrence Kohlberg's theory (Kohlberg, 1984). The DIT2 consists of five ethical dilemmas with twelve issues following each dilemma. Respondents rate and rank the issues in order of importance. These responses are analyzed to determine several scores. The primary score of interest for this study, the P score, reflects the percentage of principled moral reasoning preferred by participants. In terms of reliability, α falls between the upper .70s and lower .80s; test-retest reliability is comparable. In the Narvaez et al. (1999) study, α was .71 for the entire sample for both studies. In this study, Cronbach's α reached only .54. This was due to a more homogeneous sample in terms of the DIT2 P scores. In addition, the reliability estimate was lower since the years of formal education were restricted in this sample (Rest & Narvaez, 1998b).

As aforementioned, political ideology was measured by one self-report item that is embedded in the DIT2. This item reads, "In terms of your political views, how would you characterize yourself?" (Rest & Narvaez, 1998a)? Respondents selected one of the following responses: Very Liberal, Somewhat Liberal, Neither Liberal nor Conservative, Somewhat Conservative, or Very Conservative. Narvaez et al. (1999) reported that this approach was used instead of one that would ask respondents to respond to political issues since the ATHRI, which is comprised of politically-related items, was being used to measure the criterion variable. In addition, they reported that other researchers had used the same approach. No psychometric data have been published for this item.

Inventory of Religious Beliefs. This study used Brown and Lowe's (1951) Inventory of Religious Belief to measure religious ideology. The 15-item inventory seeks to measure the level of agreement with beliefs that reflect conservative Christianity. Items deal with issues like life after death, beliefs about Scripture, Jesus' virgin birth, salvation, and evolution. Bassett (1999) reported that the split-half reliability was .77 and that the Spearman-Brown formula yielded a coefficient of .87. In the Narvaez et al. (1999) study, Cronbach's alpha was .95. In this study, Cronbach's alpha reached .76, which may be due to the religious homogeneity of the sample. The range of possible scores is from 15, which indicates low agreement with Christian beliefs, to 75, which reflects agreement with these issues of Christian dogma. The items are measured on a Likert-type scale from 1 (strongly agree) to 5 (strongly disagree). To maintain consistency with the study being replicated, the scores were reversed so religious conservatism was indicated by higher scores.

Attitudes Toward Human Rights Inventory (ATHRI). The ATHRI (Getz, 1985) was used to measure students' moral thinking by assessing their views on public policy issues. The instrument consists of 48 items, while the version used in the Narvaez et al. (1999) study consisted of the original 40 items (Getz, 1985). To accurately replicate the Narvaez et al. study, only the 40 original items were used in this study. Each of the 40 items is scored on a five-point Likert-type scale. Item content includes questions on abortion, free speech, women's roles, euthanasia, homosexuality, religious freedom, and the role of government and limits on its authority. Scores range from 40 to 200, with higher scores indicating a leaning toward advocacy for human rights issues. On the original scale lower scores corresponded with the advocacy of civil rights; however, to maintain consistency with the Narvaez et al. (1999) study, the scores were reversed. In terms of reliability, the ATHRI had strong reliability in the Narvaez et al. (1999) study ($\alpha = .93$). In this study, Cronbach's α was .80. Again, this was likely due to the lack of considerable variance in the sample.

Procedure

Permission to conduct the research was provided through the chief student development officers and other appropriate personnel on both campuses. Lists of classes with primarily freshmen or seniors in them were requested. Once the lists were received, a systematic sampling of courses based on a distribution by disciplines and departments was conducted. Once this stage of sampling was completed, classes were randomly sampled until roughly 125 students at each school for each classification (i.e., freshman or senior) were identified. Then, the official at each school was contacted to request permission to complete the administration of the questionnaires in the identified classes. In turn, the officials contacted the instructors of the classes to seek permission. Classes were selected until at least 125 students per school per classification completed the batteries.

The researcher traveled to each campus to visit the classes. After explaining the nature of the study, the researcher provided students who agreed to participate with the informed consent form, requesting that they sign and return it, and with the three instruments to complete. The instruments were coded to ensure confidentiality and matched for each respondent. The three instruments were presented in random order to attempt to control for order effects. Once the informed consent forms and questionnaires were completed, they were returned to the researcher. No inducements were used.

Results

Means and standard deviations were calculated for the following scores: DIT2 P, the Inventory of Religious Beliefs (IRB), the political ideology item on the DIT2, and the ATHRI. Table 1 displays these descriptive statistics for the entire sample and for each institution. In addition, the results from the second study from Narvaez et al. (1999) are provided for comparative purposes. Comparing this study's descriptive results with the Narvaez et al. college sample should provide some perspective on the relative conservatism of this sample.

One-sample t tests ($df = 428$) were conducted on each of the variables for the overall sample using the Narvaez et al. (1999) means as the comparison amounts. Each of

the variables was significantly different at the $p < .001$ level. DIT2 P scores can range from 0 to 95, indicating the percentage of principled moral reasoning preferred by the individual. The entire sample for this study scored much lower than the Narvaez et al. sample, and the standard deviation was somewhat smaller for this study, reflecting the homogeneity of the sample. The IRB total variable has possible values of 15 to 75, with higher scores indicating religious conservatism. This study's sample mean score was close to the top of the range, which was significantly higher than the Narvaez et al. finding. In addition, the standard deviation was much smaller for this study. These results confirmed that this study's sample was extremely religiously conservative. The political ideology item was measured on a Likert scale from 1 to 5, with higher scores indicating a more conservative self-rating. The significant difference between the samples' political ideology scores indicated that this study's sample was much more politically conservative. Interestingly, the standard deviation scores were nearly identical. The ATHRI Totals can range from 40, which indicates a more conservative mindset toward critical social issues and less advocacy of civil liberties, to 200, which signifies a liberal stance. This study's sample scored significantly lower, signifying its conservatism toward advocacy for civil rights, plus its standard deviation is slightly smaller, showing the homogeneity of the sample again. In summary, these results indicated that the sample for this study was considerably more conservative on each measure than the sample in the comparison study.

To determine the relationships and potential multicollinearity among the variables in the multiple regression equation to be tested, Pearson product-moment correlation analyses were run on each variable pair. The coefficients are listed in Table 2. Five of the coefficients among the variables reached statistical significance. The strongest r value (i.e., $-.35$) was between the ATHRI total and the political item, indicating that only 12.3% of the variance can be explained in one variable by the other. The first study in the Narvaez et al. (1999) project found a stronger relationship with an r value of $-.58$ ($r^2 = 33.6\%$) which accounted for nearly three times the variance between the variables. This pattern of weaker correlations in this study as compared to Narvaez et al. remained consistent with each of the pairs of variables. Although a number of the correlation coefficients reached statistical significance, the multicollinearity among the variables was not at a level that compromised the results of the multiple regression analyses or the structural equation modeling (Garson, 2003; Licht, 1995; Sheskin, 2000).

A multiple regression analysis was run using the predictor variables (i.e., DIT2 P, political ideology, and religious ideology) to explain the variance in criterion variable, ATHRI scores. By conducting this analysis the R^2 values and β weights from this study could be compared to the findings in Narvaez et al. (1999). The regression model yielded a statistically significant result ($F = 39.57$, $df = 3$, $p < .001$, $R = .47$); however, the R^2 value (.22) indicated that only a small amount of the variance was explained by the predictor variables. These results indicated that the model did account for a significant amount of variance in moral thinking in conservative Christian college students. Table 3 displays the regression results in terms of B , the standard error of B , β , and t for the entire sample.

These results are of particular interest since this study sought to replicate the Narvaez et al. (1999) methods with a different population. In the second study in Narvaez et

al., which was based on the sample of students from a major Midwestern university, the political item, IRB total, and DIT2 P score predicted a significant amount of variance in the ATHRI with $R = .82$, which compared to $R = .47$ for the entire sample in this study. The β weights from that study were .27 for the DIT2 P score, $-.25$ for the IRB total, and $-.52$ for the political item. These values compared to .29, $-.11$, and $-.30$ respectively in this study. Therefore, the P score achieved a similar weight in this study, while the IRB and political items did not. These findings indicated that the P score was as strong a predictor of moral thinking in the Narvaez et al. study as in this study. However, the IRB and political variables did not account for as much variance in moral thinking in this study as in Narvaez et al.

Although previous studies that used the moral thinking prediction model did not use structural equation modeling to assess the model's fit with the data from those studies, the model lent itself to confirmatory analysis (Byrne, 2001). Another key reason for using structural equation modeling was that the cultural ideology variable, the variable comprised of the political ideology item and the IRB, could not be measured directly as an unobserved or latent variable (Arbuckle & Wothke, 1999; Byrne, 2001). Since structural equation modeling enables the researcher to present a causal model and to display the direct and indirect effects among the variables (Pedhazur, 1997), this technique was used, using the DIT2 P score and cultural ideology, comprised of the IRB total and the political item, to predict to ATHRI. The maximum likelihood for estimating the model was used. Table 4 provides the weights for the model, the standard error of the estimate, the critical ratios for the paths, and the corresponding p values. Figure 1 displays the path diagram. The diagram includes standardized regression weights since the B values were in different units of measurement, facilitating easier comparison of the "magnitude of effects of different causes" (Cohen, Cohen, West, & Aiken, 2003, p. 464) from the different variables.

To determine the overall goodness of fit of the model, a χ^2 test was run. A good model is characterized by a low χ^2 score that does not reach statistical significance (Cohen et al., 2003). The χ^2 value for the model was 5.20 ($df = 2$; $p = .074$), which did not reach statistical significance. However, Hoelter's Critical N , the size of the sample needed to accept the χ^2 results at the .05 level, was 493. Therefore, the model cannot be accepted based on the χ^2 results due to the insufficient sample size. However, Garson (2003) recommended using more than the χ^2 test as the sole determinant of goodness-of-fit. Therefore, the root mean square error of approximation (RMSEA) was used to determine the goodness-of-fit as well. RMSEA "does not require the author [to] posit as plausible a model in which there is complete independence of the latent variables" (Garson, 2003, p. 17), unlike other indicators, and is not affected much by sample size like χ^2 . A model has good fit if the RMSEA score is $\leq .05$ and adequate fit if the score is $\leq .08$. The RMSEA score for the model was .061, indicating that the model had adequate fit. In addition, certain measures "are appropriate when comparing models which have been estimated using maximum likelihood estimation" (Garson, 2003, p. 18). One such measure is the Browne-Cudeck criterion. To assume good fit, the Browne-Cudeck criterion should be close to .9. This value was 29.49, indicating a lack of fit. Since two of the measures did not indicate good fit, the model cannot be accepted. Although each of the paths in Table 4 reached significance ($p < .001$), they are meaningless since the overall model could not be accepted (Garson, 2003).

Discussion

This study indicated that the regression model does predict a significant amount of variance in moral thinking in students at evangelical Christian liberal arts institutions; however, the model does not have good statistical fit. Moreover, though the model's regression results were significant, the amount of variance predicted was much lower for this study as compared to other published studies (Narvaez et al., 1999; Rest, Narvaez, Thoma, et al., 1999). The sample for this study was very conservative religiously and politically and was less apt to advocate for civil rights as compared to the Narvaez et al. (1999) study. These differences were expected since students were sampled from evangelical colleges. However, the DIT2 P scores were significantly lower than the students from the Narvaez et al. study, who were sampled from a large Midwestern university. This was somewhat surprising since Pascarella and Terenzini (1991) found that the highest scoring type of institution was the church-affiliated liberal arts college. However, very little research on moral judgment has been done in very conservative evangelical Christian liberal arts colleges, and the campuses selected for this study were likely more conservative than those Christian liberal arts schools studied before. With this in mind, the findings from this study seemed to confirm the literature concerning moral judgment and religion which consistently points to the relationship between religious conservatism and lower postconventional thinking (Rest, Narvaez, Bebeau, et al. 1999). The moral judgment scores were likely influenced considerably by the conservative political and religious ideologies of the students. Perhaps the students had the ability to think at higher levels but chose to use faith-based principles to make moral decisions, as was the case with the fundamentalist seminarians in Lawrence's study (1979).

There were a few key limitations to this study. The sampling used limits the generalizability of the findings to the population of all students at Christian colleges. The multistage sampling procedure presents several key problems. The schools sampled are in the Southeastern United States, while the vast majority of CCCU member institutions are outside of this region. In addition, each school is associated with a different denomination or faith tradition which, in turn, influences the schools and their students in different ways (e.g., how religion and ethics are taught, how students are exposed to particular social and political commitments, etc.). The research design for this study does not account for these differences which may influence student responses. Therefore, generalizing to all CCCU members or Christian colleges may be questionable. In addition, as discussed by Pascarella and Terenzini (1991), the observational methods used in this study will not allow for definitively answering the question of whether any of the changes in moral development can be attributed to a specific college effect or maturation. Specific to this study would be the difficulty in substantiating claims that Christian colleges "caused" certain effects. Furthermore, the range of responses on the instruments used in this study was restricted due to the homogeneity of the sample. This resulted in attenuated coefficients in correlational and regression analyses. In addition, it likely decreased the reliability estimates of the instruments.

The primary implication of this study is that evangelical Christian liberal arts colleges, which accent student moral development, can use the model to help them predict how their students think about significant social and political issues. Having

such models should help such schools assess their students' moral development outcomes, thereby demonstrating that they have accomplished their missions. This is of particular importance since even schools regarded as having exemplary moral and civic development programs seldom assess these outcomes. Historically, schools have chosen not to assess these programs and have lacked valid and reliable tools to do so. Assessment models, such as the one utilized in this study, can help these campuses assess their mission achievement, improve in these areas, and inform their programs (Colby, Ehrlich, Beaumont, & Stephens, 2003). As schools begin to use results from assessment models like this, they can determine or tailor specific interventions that can facilitate the desired change. Obviously, this is predicated on the idea that colleges have a sense of what moral thinking they desire in their students.

However, schools must be mindful that the model accounted for a very low amount of variance in moral thinking and lacked good statistical fit. With this in mind, these institutions must assess the fit of the model on their campuses, and when indicated, include other predictor variables consistent with the literature to enhance the model's fit. In fact, schools can develop specific measures for themselves to include in the model. These measures would be particularly useful if there are specific programs that encourage moral discourse and reflection. Some recent research by McNeel, Frederickson, and Granstrom (1998) has enhanced the model's predictive power with a more religiously conservative sample than in the Narvaez et al. (1999) study by adding measures of how participants hold their faith. In essence, these measures assessed whether conservative Christians approached their faith dogmatically or were open to other insights to their faith. Christians who held their faith less dogmatically tended to endorse positions that were more supportive of human rights. Perhaps these or similar measures should be used when using the model with conservative Christians. In addition, these models should be assessed for goodness-of-fit.

Other fruitful areas for research include using cognitive ability as a predictor since it correlates highly with moral judgment yet is distinct from it (Rest, 1979). In addition, other recent research has shown that growth in moral reasoning was enhanced by a college's curriculum and the student's ability to think critically (Mentkowski & Associates, 2000). The link between critical thinking and moral reasoning was more pronounced in the first two years of college. Therefore, further research should evaluate the role of critical thinking in predicting attitudes toward human rights. Furthermore, certain aspects of the institution's culture or ethos could be assessed, especially since the "hidden curriculum" tends to have a strong influence on morality (Colby et al., 2003). For instance, the level of academic challenge at an institution may affect the level of critical thinking achieved by students which, in turn, may affect the level of moral judgment. Obviously, some of the institutional characteristics, the campus culture itself, and student subcultures could be assessed more thoroughly through qualitative methods like interviews, document analysis, focus groups, and observation (Kuh & Whitt, 1988; Upcraft & Schuh, 1996; Whitt, 1996). By doing this, these studies could understand more fully how the college affected students' moral thinking. One particular issue related to the institution's effects on moral thinking that should be considered in future studies is the degree to which moral development is central to the mission and goals of the college. As Colby et al. (2003) identified in their study of schools that promoted moral and civic development, "Leadership from administrators,

faculty, and campus centers is central to their success, as is establishing a campus culture that supports positive moral and civic values" (p. xv). For schools to facilitate student moral development, they must address these issues in the core and major curricula and offer experiences outside of the classroom that contribute to this growth.

References

- Arbuckle, J. L., & Wothke, W. (1999). *Amos 4.0 user's guide*. Chicago, IL: SmallWaters Corporation.
- Babbie, E. (1990). *Survey research methods* (2nd ed.). Belmont, CA: Wadsworth.
- Bassett, R. (1999). Review of Inventory of Religious Belief. In P. C. Hill & R. W. Hood, Jr. (Eds.), *Measures of religiosity* (pp. 22-24). Birmingham, AL: Religious Education Press.
- Baylis, B. (1997). Report on the 1996 first-year, entering students. *Research on Christian Higher Education*, 4, 107-117.
- Beller, J. M., Stoll, S. K., Burwell, B., & Cole, J. (1996). The relationship of competition and a Christian liberal arts education on moral reasoning of college student athletes. *Research on Christian Higher Education*, 3, 99-114.
- Brown, D. E., & Lowe, W. L. (1951). Religious beliefs and personality characteristics of college students. *Journal of Social Psychology*, 33, 103-29.
- Brubacher, J. S., & Rudy, W. (1997). *Higher education in transition: A history of American colleges and universities* (4th ed.). New Brunswick, NJ: Transaction.
- Burwell, R. J. (1997). Report on the 1996 senior College Student Survey. *Research on Christian Higher Education*, 4, 119-130.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Colby, A., Ehrlich, T., Beaumont, E., & Stephens, J. (2003). *Educating citizens: Preparing America's undergraduates for lives of moral and civic responsibility*. San Francisco: Jossey-Bass.
- Council for Christian College and Universities. (2000). *CCCU directory and resource guide for Christian higher education 2000-2001*. Washington, DC: Author.
- Dirks, D. H. (1988). Moral development in Christian higher education. *Journal of Psychology and Theology*, 16, 324-331.
- Dobson, J. C. (1998). Our perspective: The benefits of Christian higher education. In Peterson's (Ed.), *Christian colleges and universities: Top institutions committed to academic quality and character development* (p. 3). Princeton, NJ: Peterson's.
- Evans, N. J., Forney, D. S., & Guido-DiBrito, F. (1998). *Student development in college: Theory, research, and practice*. San Francisco: Jossey-Bass.
- Fowler, F. J., Jr. (1993). *Survey research methods* (2nd ed.). Newbury Park, CA: Sage.
- Garber, S. (1996). *The fabric of faithfulness: Weaving together belief and behavior during the university years*. Downers Grove, IL: Inter Varsity Press.
- Garson, G. D. (n.d.). *Structural equation modeling*. Retrieved October 17, 2003, from North Carolina State University Web site: <http://www2.chass.ncsu.edu/garson/pa765/structur.htm>
- Getz, I. R. (1984). Moral judgment and religion: A review of the literature. *Counseling and Values*, 28, 94-116.
- Getz, I. R. (1985). *Moral reasoning, religion, and attitudes toward human rights*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.
- Henry, G. T. (1990). *Practical sampling*. Newbury Park, CA: Sage.
- Homles, A. F. (1987). *The idea of a Christian college* (rev. ed.). Grand Rapids, MI: Eerdmans.
- Holmes, A. F. (1991). *Shaping character: Moral education in the Christian college*. Grand Rapids, MI: Eerdmans.
- Kerr, C. (1995). *The uses of the university* (4th ed.). Cambridge, MA: Harvard.
- Kohlberg, L. (1981). *Essays on moral development: The philosophy of moral development: Moral stages and the idea of justice* (Vol. 1). San Francisco: Harper & Row.
- Kohlberg, L. (1984). *Essays on moral development: The psychology of moral development: The nature and validity of moral stages* (Vol. 2). San Francisco: Harper & Row.
- Kuh, G. D., & Whitt, E. J. (1988). *The invisible tapestry: Culture in American colleges and universities*. ASHE-ERIC Higher Education Report No. 1. Washington, DC: Association for the Study of Higher Education.
- Lawrence, J. A. (1979). The component procedures of moral judgment making (Doctoral dissertation, University of Minnesota, Minneapolis, 1979). *Dissertation Abstracts International*, 40, 896B.
- Licht, M. H. (1995). Multiple regression and correlation. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 19-64). Washington, DC: American Psychological Association.
- McNeel, S. P. (1991). Christian liberal arts education and growth in moral judgment. *Journal of Psychology and Christianity*, 10, 311-322.
- McNeel, S. P. (1992). *Moral maturing in college*. Unpublished manuscript, Bethel College at St. Paul.
- McNeel, S. P. (1994). College teaching and student moral development. In J. Rest & D. Narvaez (Eds.), *Moral development in the professions: Psychology and applied ethics* (pp. 27-50). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McNeel, S. P., Frederickson, J., & Granstrom, S. (1998, November). *Attitudes toward human rights: Moral, religious and political predictors*. Paper presented at the meeting of the Association of Moral Education, Durham, NH.
- Mentkowski, M., & Associates. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco: Jossey-Bass.
- Narvaez, D., Getz, I., Rest, J. R., & Thoma, S. J. (1999). Individual moral judgment and cultural ideologies. *Developmental Psychology*, 35, 478-488.
- Nucci, L., & Pascarella, E. T. (1987). The influence of college on moral development. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research volume III* (pp. 271-326). New York: Agathon Press, Inc.

- Pascarella, E. T. (1997, Summer/Fall). College's influence on principled moral reasoning. *Educational Record*, 47-55.
- Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass Publishers.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace.
- Peterson's. (1998). *Christian colleges and universities: Top institutions committed to academic quality and character development*. Princeton, NJ: Author.
- Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis: University of Minnesota Press.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.
- Rest, J. R., & Narvaez, D. (1998a). *Defining Issues Test – Second Edition*. Minneapolis, MN: Center for Research in Ethical Development.
- Rest, J. R., & Narvaez, D. (1998b). *Guide for DIT-2*. Minneapolis, MN: Center for the Study of Ethical Development.
- Rest, J. R., Narvaez, D. F., Bebeau, M. J., & Thoma, S. J. (1999). *Postconventional moral thinking: A neo-Kohlbergian approach*. Hillsdale, NJ: Lawrence Erlbaum.
- Rest, J. R., Narvaez, D. F., Thoma, S. J., & Bebeau, M. J. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology*, 91, 644-659.
- Sheskin, D. J. (2000). *Handbook of parametric and nonparametric statistical procedures* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Upcraft, M. L., & Schuh, J. H. (1996). *Assessment in student affairs: A guide for practitioners*. San Francisco: Jossey-Bass.
- Whitt, E. J. (1996). Assessing student cultures. In M. L. Upcraft, & J. H. Schuh (Eds.), *Assessment in student affairs: A guide for practitioners* (pp. 189-216). San Francisco: Jossey-Bass.

Table 1.

Means and standard deviations for moral judgment, religious ideology, political ideology, and attitudes toward human rights

| Variable | Epsilon | Theta | All | Narvaez ^a | t ^b |
|-----------|-------------------|-------------------|-------------------|----------------------|----------------|
| DIT2 P | 33.36 (14.47) | 28.60 (12.60) | 30.81 (13.69) | 48.58 (15.13) | 19.40 *** |
| IRB | 70.18 (4.92) | 70.33 (4.76) | 70.26 (4.83) | 55.48 (14.78) | -35.91 *** |
| Political | 3.79 (0.92) | 3.65 (0.92) | 3.71 (0.92) | 2.85 (0.94) | -26.88 *** |
| ATHRI | 136.77 (13.76) | 136.12 (12.56) | 136.42 (13.12) | 159.16 (17.26) | 63.44 *** |

Note. Standard deviations are in parentheses. DIT2 P = Defining Issues Test 2 P score; IRB = Inventory of Religious Beliefs; Political = political ideology item; ATHRI = Attitudes Towards Human Rights Inventory.

^a Narvaez et al. (1999) Study II

^b t test difference is the one-sample t test for differences between the entire sample for this study and the sample for the second study in Narvaez et al. (1999).

*** p < .001.

Table 2.

Correlations between variables

| Variable | DIT2 P | Political | IRB | ATHRI |
|-----------|---------|-----------|---------|-------|
| DIT2 P | --- | | | |
| Political | -.06 | --- | | |
| IRB | -.11 * | .31 ** | --- | |
| ATHRI | .31 *** | -.35 ** | -.23 ** | --- |

Note. DIT2 P = Defining Issues Test 2 P score; Political = political ideology; IRB = Inventory of Religious Beliefs; ATHRI = Attitudes Towards Human Rights Inventory.

* p < .05. ** p < .01. *** p < .001

Table 3.

Multiple regression results

| Variable | B | SE B | β | t | Sig. |
|-----------|-------|------|---------|-------|------|
| P score | .29 | .04 | .27 | 6.60 | *** |
| Political | -4.27 | .64 | -.30 | -6.63 | *** |
| IRB | -.29 | .12 | -.11 | -2.32 | * |

Note. DIT2 P = Defining Issues Test 2 P score; IRB = Inventory of Religious Beliefs;

Political = political ideology.

* $p < .05$. *** $p < .001$.**Table 4.**

Regression weights for Regression Model

| Path | Estimate | SE | CR | <i>p</i> |
|---------------------------------|----------|------|-------|----------|
| ATHRI \leftarrow P score | 0.27 | .04 | 6.66 | .000 |
| IRB \leftarrow Cultural | 3.14 | .72 | 4.38 | .000 |
| Political \leftarrow Cultural | 1.00 | | | |
| ATHRI \leftarrow Cultural | -9.03 | 2.06 | -4.39 | .000 |

Note. DIT2 P = Defining Issues Test 2 P score; IRB = Inventory of Religious Beliefs;

Political = political ideology; ATHRI = Attitudes Toward Human Right Inventory.

Figure 1. Path diagram for predicting moral thinking.

Note. P score = Defining Issues Test 2 P score; IRB = Inventory of Religious Beliefs; Political = political ideology; ATHRI = Attitudes Toward Human Right Inventory.

